

OM nucleic - nucleic search, using sw model
Run on: December 5, 2003, 07:48:44 ; Search time 2892 Seconds
(without alignments)
10439.594 Million cell updates/sec
Title: US-09-786-715-5
Perfect score: 738
Sequence: 1 gcaccaggaattttagt.....aaaaaaaaaaaaa 738
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdt:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	289.6	39.2	630	8	PSA310990	AJ310990 Pisum sat
2	233.6	31.7	603	8	KCH10RNN	Z70677 R.communis
3	211.6	28.7	665	8	AV170650	AY170650 Pisum sat
4	203.4	27.6	698	8	NTRNA	X38527 N.tabacum m
5	201.8	27.3	480	8	AV088687	AY088687 Arabidops
6	201.8	27.3	497	8	ATH10ARA	Z14084 A.thaliana
7	198.6	26.9	345	6	AX505468	AX505468 Sequence
8	192	26.0	653	6	AR016869	AR016869 Sequence
9	192	26.0	653	6	AR020895	AR020895 Sequence
10	192	26.0	653	6	AR027218	AR027218 Sequence
11	192	26.0	653	6	AR038505	AR038505 Sequence
12	192	26.0	653	6	AR064647	AR064647 Sequence
13	192	26.0	653	6	AR067572	AR067572 Sequence
14	192	26.0	653	6	I38524	I38524 Sequence 10
15	192	26.0	653	6	I56999	I56999 Sequence 10
16	192	26.0	653	6	I59865	I59865 Sequence 10
17	192	26.0	653	6	I75192	I75192 Sequence 10
18	188	25.5	664	8	AF323593	AF323593 Prunus pe
19	186.6	25.3	784	8	AY271308	AY271308 Citrus x
20	182.2	24.7	593	8	PSA319808	AJ319808 Pisum sat
21	181	24.5	345	8	AF483265	AF483265 Populus t
22	173.2	23.5	580	8	BNUS8380	U59380 Brassica na
23	172.8	23.4	590	8	AY088698	AY088698 Arabidops
24	171.2	23.2	556	8	AK118542	AK118542 Arabidops
25	170.6	23.1	360	6	AX505469	AX505469 Sequence
26	170.6	23.1	360	8	BT004710	BT004710 Arabidops
27	170.4	23.1	528	8	ATH1RED4	Z35476 A.thaliana
28	170.4	23.1	530	8	AK118035	AK118035 Arabidops
29	170.4	23.1	560	8	AY087159	AY087159 Arabidops
30	170.4	23.1	561	8	AF360227	AF360227 Arabidops
31	168.8	22.9	642	8	ATH1RED1	Z35473 A.thaliana
32	168.6	22.8	357	8	AY040028	AY040028 Arabidops
33	158.8	21.5	740	8	AB053294	AB053294 Oryza sat

34 156.2 21.4 639 8 AV085117 Arabidops
35 156.6 21.2 357 6 AX412276 Sequence
36 156.6 21.2 357 6 AX412277 Sequence
37 156.6 21.2 357 6 AX412278 Sequence
38 156.6 21.2 357 6 AX412279 Sequence
39 156.6 21.2 357 6 AX412473 Sequence
40 156.6 21.2 357 6 AX412474 Sequence
41 156.6 21.2 357 6 AX412566 Sequence
42 156.6 21.2 357 6 AX412917 Sequence
43 156.6 21.2 357 6 AX507041 Sequence
44 156.6 21.2 357 6 AX651357 Sequence
45 156.6 21.2 402 8 AX114566 Arabidops

Search completed: December 5, 2003, 09:53:26
Job time : 2898 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 06:08:34; Search time 261 Seconds
(without alignments)
7089.627 Million cell updates/sec

Title: US-09-786-715-5

Perfect score: 738

Sequence: 1 gcaccaggaaatttttagt.....aaaaaaaaaaaaaaaa 738

Scoring table: IDENTITY NJC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	738	100.0	738	21	AAZ51739	Glycine max thior
2	208.8	28.3	402	25	ABX21664	Human GDP-mannose
3	201.8	27.3	561	21	AAC31522	Arabidopsis thalia
4	201.8	27.3	563	21	AAC34121	Arabidopsis thalia
5	201.6	27.3	3129	24	ABX53095	DNA encoding thior
6	201.6	27.3	3129	24	ABN99579	Phaseolin promoter
7	201.6	27.3	3888	24	ABX53096	DNA encoding Oleos
8	201.6	27.3	3888	24	ABN99580	Phaseolin promoter
9	199.2	27.0	4935	24	ABN99586	Promoter-Oleosin t
10	198.6	26.9	345	24	ABZ12358	Arabidopsis thalia
11	198.6	26.9	345	24	ABN99587	Arabidopsis thalia
12	197.4	26.7	3888	24	ABX53097	DNA encoding Thior
13	197.4	26.7	3888	24	ABN99581	Phaseolin promoter
14	195.6	26.5	470	21	AAC37781	Arabidopsis thalia
15	193.4	26.2	392	23	ABX16045	Human GDP-mannose
16	192	26.0	653	16	AAQ99753	Plant SAR gene pl.
17	192	26.0	653	20	AAV62799	Tobacco SAR CHX in
18	192	26.0	653	20	AAV81683	Tobacco protein-sy
19	191.6	26.0	574	21	AAZ51738	Catalpa speciosa t
20	186.2	25.2	601	21	AAZ51740	Glycine max thior
21	176.6	23.9	390	25	ABX19403	Human GDP-mannose
22	175.6	23.8	328	21	AAZ31097	Plant microsatelli
23	172.8	23.4	509	22	AAH87768	Peppermint plant o
24	172.8	23.4	590	21	AAC38792	Arabidopsis thalia
25	170.6	23.1	360	24	ABZ12359	Arabidopsis thalia
26	170.4	23.1	560	21	AAZ41961	Arabidopsis thalia
27	170.2	23.1	614	21	AAZ51741	Vernonia mespilifo
28	169.8	23.0	419	21	AAZ31942	Plant microsatelli
29	168.6	22.8	346	21	AAZ31096	Plant microsatelli
30	167.6	22.7	320	21	AAZ31785	Plant microsatelli
31	163	22.1	292	21	AAZ31978	Plant microsatelli
32	158.8	21.5	356	21	AAZ31804	Plant microsatelli
33	158.2	21.4	480	21	AAC36542	Arabidopsis thalia
34	157.2	21.3	353	21	AAZ31118	Plant microsatelli
35	156.8	21.2	524	21	AAC33829	Arabidopsis thalia
36	156.6	21.2	357	24	ABZ13931	Arabidopsis thalia
37	156.6	21.2	652	21	AAC45656	Arabidopsis thalia
38	155.6	21.1	277	25	ABX26076	Human GDP-mannose
39	155.2	21.0	572	21	AAC52069	Arabidopsis thalia
40	155	21.0	563	21	AAC34211	Arabidopsis thalia
41	153.6	20.8	870	21	AAZ51737	Memordica charanti
42	152.8	20.7	328	21	AAZ31222	Plant microsatelli
43	152.2	20.6	274	21	AAZ31859	Plant microsatelli
44	151.8	20.6	576	25	ABX56868	Arabidopsis thalia
45	149.4	20.2	647	21	AAC38520	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAZ51739	
ID	AAZ51739 standard; cDNA; 738 BP.
XX	
AC	AAZ51739;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	Glycine max thior
XX	
KW	Glycine max thior; clone sahlc-pk001.117; chimeric gene; soybean;
KW	transgenic plant; seed storage protein; allergenicity; ss.
XX	
OS	Glycine max.
XX	
FR	Key Location/Qualifiers
FT	CDS 68..430
FT	/*tag= a
FT	/product= "Thior
XX	
PN	W0200014239-A2.
XX	
PD	16-MAR-2000.
XX	
PF	07-SEP-1999; 99WO-US20420.
XX	
PR	08-SEP-1998; 98US-0099501.
XX	
PA	(DUPO) DU FONT DE NEMOURS & CO E I.
XX	
PI	Allen SM, Thorpe CU, Lu AL;
XX	
DR	WPI; 2000-256987/22.
DR	P-PSDB; AA70481.
XX	
PT	New isolated polynucleotide encoding thior polypeptide is useful
PT	for producing transgenic plants with an altered level of thior:
XX	
PS	Claim 3; Page 30; 33pp; English.
XX	
CC	The present cDNA sequence encodes Glycine max thior protein. The
CC	cDNA was derived from clone sahlc-pk001.117, which was isolated from a
CC	cDNA library prepared from soybean tissue sprayed with authority
CC	herbicide. Chimeric genes encoding all or a portion of the thior
CC	protein, in sense or antisense orientation are constructed, wherein
CC	expression of the chimeric gene results in production of altered levels
CC	of the thior protein in a transformed host cell. Thior is
CC	involved in the disassembly of seed storage proteins during germination
CC	by reducing S-S bonds and in the bread making process. Over expression of
CC	thior in cereals may reduce the allergenicity of any transgenic
CC	protein engineered into cereal crops with high sulphydryl content.
XX	
SQ	Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;

Query Match 100.0%; Score 738; DB 21; Length 738;
Best Local Similarity 100.0%; Pred. No. 3e-101;

Matches 739; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCACAGGAATTCCTTAGTTGTAACCTGACAAAGTTTCTGAGAAATAGAGATTATAT	60
Db	1	GCACAGGAATTCCTTAGTTGTAACCTGACAAAGTTTCTGAGAAATAGAGATTATAT	60
Qy	61	TGAGAGATGGCTGGCTCATCGGAAGGAGGCAAGTCATTAGCTGCCACACCGTTGAAGA	120
Db	61	TGAGAGATGGCTGGCTCATCGGAAGGAGGCAAGTCATTAGCTGCCACACCGTTGAAGA	120
Qy	121	ATGGAACGATCACTCCAGAGGGCAACGATCCAGAACTCATTTGTGTGGATTAC	180
Db	121	ATGGAACGATCACTCCAGAGGGCAACGATCCAGAACTCATTTGTGTGGATTAC	180
Qy	181	TGCTTCTTGGGTGGACCATGCGTTTTCATGACCATCTCTGGCTGAGCTGCTAAAGA	240
Db	181	TGCTTCTTGGGTGGACCATGCGTTTTCATGACCATCTCTGGCTGAGCTGCTAAAGA	240
Qy	241	GTTCCACAGTGTCTATTCCTTAAAGGTGGATGGAGCAATTAAGAGTGTTCCTCAAGA	300
Db	241	GTTCCACAGTGTCTATTCCTTAAAGGTGGATGGAGCAATTAAGAGTGTTCCTCAAGA	300
Qy	301	TGGGCTATGAGGCTATGCCACTTTTGTGTGTGTAAGAGGAGCGTTCTGGACA	360
Db	301	TGGGCTATGAGGCTATGCCACTTTTGTGTGTGTAAGAGGAGCGTTCTGGACA	360
Qy	361	AGTGGTGGGCAAGAGGATGAGCTGCAGCAGAAATACAGAACTGTGGCTTCAGC	420
Db	361	AGTGGTGGGCAAGAGGATGAGCTGCAGCAGAAATACAGAACTGTGGCTTCAGC	420
Qy	421	TAGTGTCTTAACTAGCTCACCTTCAGAACTTTATATAGGCTTTCTTTTCAATCT	480
Db	421	TAGTGTCTTAACTAGCTCACCTTCAGAACTTTATATAGGCTTTCTTTTCAATCT	480
Qy	481	GTACTAGACTTATGTTGGTATTTCTGTATTGCACCATCAGCTTTCAAAGGTGATGAC	540
Db	481	GTACTAGACTTATGTTGGTATTTCTGTATTGCACCATCAGCTTTCAAAGGTGATGAC	540
Qy	541	TCCTATCATCTATTCCTGAATAGTAGTAAGTGGTCTTTCTTCGCTCTTAATAATAGT	600
Db	541	TCCTATCATCTATTCCTGAATAGTAGTAAGTGGTCTTTCTTCGCTCTTAATAATAGT	600
Qy	601	GATGGTGTATATCAATGATCTTAATACATAGACCTTCTGTCTTCCTTTTAGTATTA	660
Db	601	GATGGTGTATATCAATGATCTTAATACATAGACCTTCTGTCTTCCTTTTAGTATTA	660
Qy	661	AA	720
Db	661	AA	720
Qy	721	AAAAAAAAAAAAAAAAAAAAA 738	
Db	721	AAAAAAAAAAAAAAAAAAAAA 738	

Search completed: December 5, 2003, 09:04:57
Job time : 284 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 09:53:30 ; Search time 318 Seconds
(without alignments)
7713.259 Million cell updates/sec

Title: US-09-786-715-5
Perfect score: 738
Sequence: 1 gcaaccaggaatttttagt.....aaaaaaaaaaaaaaaaaaaa 739

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1 208.8 28.3 402 10 US-09-878-574-3723
2 201.6 27.3 3129 9 US-09-897-898-5
3 201.6 27.3 3129 12 US-10-032-201B-14
4 201.6 27.3 3888 9 US-09-897-898-7
5 201.6 27.3 3888 12 US-10-032-201B-16
6 198.2 27.0 4935 12 US-10-032-201B-33
7 198.6 26.9 345 10 US-09-938-842A-163
8 198.6 26.9 345 12 US-10-032-201B-38
9 197.4 26.7 3888 9 US-09-897-898-10
10 197.4 26.7 3888 12 US-10-032-201B-19
11 193.4 26.2 392 10 US-09-878-574-104
12 176.6 23.9 390 10 US-09-878-574-1462
13 170.6 23.1 360 10 US-09-938-842A-164
14 156.6 21.2 357 10 US-09-938-842A-1736
15 156.6 21.2 357 14 US-10-323-362-1
16 155.6 21.1 277 10 US-09-878-574-8135
17 144.4 19.6 370 10 US-09-878-574-3169
18 140.4 19.0 268 10 US-09-878-574-5995
19 140.2 19.0 540 12 US-10-349-782-12
20 139.2 18.9 382 12 US-10-091-841-3
21 139.2 18.9 393 12 US-10-091-841-5
22 138.2 18.7 260 10 US-09-878-574-12941
23 135.6 18.4 369 12 US-10-394-865-8
24 135.6 18.4 369 12 US-10-091-841-1
25 135 18.3 270 10 US-09-878-574-13445
26 130.8 17.7 272 10 US-09-878-574-12181
27 129.8 17.6 659 12 US-10-349-782-9
28 120 16.3 267 10 US-09-878-574-13137
29 105.2 14.3 402 10 US-09-938-842A-225
30 105.2 14.3 613 9 US-09-770-149-855
31 103.8 14.1 456 10 US-09-924-035A-677
32 100.6 13.6 297 12 US-10-349-782-11
33 94.6 12.8 250 9 US-09-923-876-1288
34 93.2 12.6 421 14 US-10-198-846-6007
35 92.2 12.5 476 12 US-09-814-353-4700
36 92.2 12.5 476 12 US-09-814-353-10999
37 88.4 12.0 470 11 US-09-918-995-22375
38 88.2 12.0 299 12 US-09-814-353-4844
39 88.2 12.0 299 12 US-09-814-353-11141
40 88 11.9 261 10 US-09-878-574-7974
41 87.2 11.8 419 10 US-09-960-352-11234
42 86.4 11.7 839 14 US-10-198-846-5829
43 86 11.7 2964 12 US-10-237-496-109
44 86 11.7 2964 12 US-10-242-074-109
45 86 11.7 2964 12 US-10-242-505-109

Search completed: December 5, 2003, 11:49:06
Job time : 321 sec

Sequence 3723, Ap
Sequence 5, Appli
Sequence 14, Appl
Sequence 7, Appli
Sequence 16, Appl
Sequence 33, Appl
Sequence 163, Appl
Sequence 38, Appl
Sequence 10, Appl
Sequence 19, Appl
Sequence 144, Appl
Sequence 1462, Ap
Sequence 164, App
Sequence 1736, Ap
Sequence 1, Appli
Sequence 8135, Ap
Sequence 3169, Ap
Sequence 5995, Ap
Sequence 12, Appl
Sequence 3, Appli
Sequence 5, Appli
Sequence 12941, A
Sequence 8, Appli
Sequence 13445, A
Sequence 12181, A
Sequence 9, Appli
Sequence 13137, A
Sequence 225, App
Sequence 855, App
Sequence 677, App
Sequence 1288, Ap
Sequence 6007, Ap
Sequence 4700, Ap
Sequence 10999, A
Sequence 2375, A
Sequence 4844, Ap
Sequence 11141, A
Sequence 7974, Ap
Sequence 11234, A
Sequence 5829, Ap
Sequence 109, App
Sequence 109, App
Sequence 109, App

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 10:29:10 ; Search time 1939 Seconds
(without alignments)
2531.797 Million cell updates/sec

Title: US-09-786-715-6

Perfect score: 615

Sequence: 1 MAGSSEGVVISCHVEMN.....GAKKDELQKQIKHVASASA 120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2688711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp

-C=/cgr2.1/USPTO.spool.p/US09786715/runat_04122003_092646_3937/app.query.fasta_1.263

-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -ILOOPEXT=0

-UNIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US09786715@cgr2.1.3745 @runat_04122003_092646_3937 -NCPU=6 -ICPU=3

-NO.MMAP -LARGEOUT=1 -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LNGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: +
1: gb_ba: +
2: gb_hg: +
3: gb_in: +
4: gb_om: +
5: gb_ov: +
6: gb_pat: +
7: gb_ph: +
8: gb_pl: +
9: gb_pr: +
10: gb_ro: +
11: gb_scs: +
12: gb_sy: +

13: gb_uni.*
 14: gb_vi.*
 15: em_pa.*
 16: em_fun.*
 17: em_hum.*
 18: em_in.*
 19: em_mu.*
 20: em_cm.*
 21: em_or.*
 22: em_ov.*
 23: em_pat.*
 24: em_ph.*
 25: em_pl.*
 26: em_rc.*
 27: em_ats.*
 28: em_un.*
 29: em_vi.*
 30: em_htg_hum.*
 31: em_htg_inv.*
 32: em_htg_mus.*
 33: em_htg_other.*
 34: em_htg_pin.*
 35: em_htg_rnd.*
 36: em_htg_man.*
 37: em_htg_vrt.*
 38: em_sy.*
 39: em_htgo_hum.*
 40: em_htgo_mus.*
 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	544	88.5	630	8	PSA310990	AJ310990 Pisum sat
2	490	79.7	603	8	RCTHICRXN	Z70677 R. communis
3	482	78.4	345	6	AX505468	AX505468 Sequence
4	482	78.4	480	8	AY088687	AY088687 Arabidops
5	482	78.4	497	8	ATTHIOARA	Z14084 A.thaliana
6	460	74.8	664	8	AF323593	AF323593 Prunus pe
7	456	74.1	698	8	NTTRNA	X58527 N.tabacum m
8	453	73.7	665	8	AY170650	AY170650 Pisum sat
9	445	72.4	653	6	AR016869	AR016869 Sequence
10	445	72.4	653	6	AR020895	AR020895 Sequence
11	445	72.4	653	6	AR027218	AR027218 Sequence
12	445	72.4	653	6	AR038505	AR038505 Sequence
13	445	72.4	653	6	AR064647	AR064647 Sequence
14	445	72.4	653	6	AR067572	AR067572 Sequence
15	445	72.4	653	6	I38524	I38524 Sequence 10
16	445	72.4	653	6	I56999	I56999 Sequence 10
17	445	72.4	653	6	I59865	I59865 Sequence 10
18	445	72.4	653	6	I75192	I75192 Sequence 10

19	434	70.6	784	8	AY271308	AY271308 Citrus x
20	429	69.8	357	8	AY040028	AY040028 Arabidops
21	429	69.8	528	8	ATTIHRED4	Z35476 A.thaliana
22	429	69.8	530	8	AK118035	AK118035 Arabidops
23	429	69.8	560	8	AY087159	AY087159 Arabidops
24	429	69.8	561	8	AF360227	AF360227 Arabidops
25	414	67.3	593	8	PSA319808	PSA319808 Pisum sat
26	399	64.9	345	8	AF483265	AF483265 Populus t
27	388	63.1	526	11	G73679	G73679 RZ488R etio
28	388	63.1	601	8	OSU92541	U92541 Oryza sativ
29	388	63.1	686	6	R08194	E08194 RPS13 gene
30	388	63.1	687	8	RICTH	I21836 Oryza sativ
31	385.5	62.7	590	8	AY088698	AY088698 Arabidops
32	385	62.6	595	8	D87984	D87984 Fagopyrum e
33	383	62.3	366	6	AX654096	AX654096 Sequence
34	381.5	62.0	360	6	AX505469	AX505469 Sequence
35	381.5	62.0	360	6	BT004710	BT004710 Arabidops
36	381.5	62.0	556	8	AK118542	AK118542 Arabidops
37	381	62.0	740	8	AE053294	AE053294 Oryza sat
38	378.5	61.5	642	8	ATTIHRED1	Z35473 A.thaliana
39	376.5	61.2	357	6	AX412276	AX412276 Sequence
40	376.5	61.2	357	6	AX412277	AX412277 Sequence
41	376.5	61.2	357	6	AX412278	AX412278 Sequence
42	376.5	61.2	357	6	AX412279	AX412279 Sequence
43	376.5	61.2	357	6	AX412473	AX412473 Sequence
44	376.5	61.2	357	6	AX412474	AX412474 Sequence
45	376.5	61.2	357	6	AX412566	AX412566 Sequence

ALIGNMENTS

RESULT 1	PSA310990	Pisum sativum mRNA for thioedoxin h (trxn gene).	630 bp	mRNA	linear	PLN 12-APR-2001
LOCUS	PSA310990	Pisum sativum mRNA for thioedoxin h (trxn gene).	630 bp	mRNA	linear	PLN 12-APR-2001
DEFINITION	AJ310990	thioedoxin H; trxn gene.				
ACCESSION	AJ310990.1	GI:13624883				
VERSION	AJ310990.1	GI:13624883				
KEYWORDS	Pisum sativum (pea)					
SOURCE	Pisum sativum					
ORGANISM	Pisum sativum					
REFERENCE	1	Traverso, J.A., Sahrawy, M., Lopez-Gorge, J. and Chueca, A. Sequencing, cloning and expression of Pisum sativum Thioedoxin h				
AUTHORS	Traverso, J.A.					
JOURNAL	Thesis (2001) Department of Bioquímica, Biología Celular y Molecular de Plantas, Estación Experimental del Zaidin (C.S.I.C.), Granada, Spain					
REFERENCE	2	Traverso, J.A., Cazalis, R., Sahrawy, M., Lopez-Gorge, J. and Chueca, A. Sequencing, cloning and expression of Pisum sativum Thioedoxin h				
AUTHORS	Traverso, J.A.					
JOURNAL	Unpublished					
REFERENCE	3	(bases 1 to 630)				
AUTHORS	Traverso, J.A.					
JOURNAL	Submitted (11-APR-2001) Traverso J.A., Bioquímica Biología Celular					

y Molecular de plantas, Estacion Experimental del Zaidin (CSIC),
Profesor Albarada i, Granada, 18008, SPAIN

FEATURES
source

..630
/organism="Pisum sativum"
/mol_type="mRNA"
/variety="Lincoln"
/db_xref="taxon:3888"
/tissue_type="Leaves"
1..630
/gene="trxh"
36..398
/gene="trxh"
/function="Unknown"
/codon_start=1
/evidence="experimental"
/product="thioredoxin h"
/protein_id="CAC36986.1"
/db_xref="GI:1362484"
/translation="MAGSSEGVVSCSHVDAMNDILHRGNSKKLIVVDFTSWGP
CRFIAPLGLGLAKFTNVIKVDVDELKVAQDWAVEAMFTFVKEGTILGWGA
KRELQQTIERHVASNA"
398..609
/gene="trxh"
609
/gene="trxh"

3'UTR

polyA_site

BASE COUNT 195 a 95 c 141 g 199 t
ORIGIN

Alignment Scores:

Pred. No.: 4.38e-55 Length: 630
Score: 544.00 Matches: 102
Percent Similarity: 95.00% Conservative: 12
Best Local Similarity: 85.00% Mismatches: 6
Query Match: 88.46% Indels: 0
DB: 8 Gaps: 0

US-09-786-715-6 (1-120) x PSA310890 (1-630)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTTPAn 20
DB 36 ATGCGAGGTTCATCAGAGAGGAGACAGTTATCAGCTGTACAGCGTTGATGCATGGAC 95
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 96 GATATCCTTCACAGAGGCAATGAATCCAGAAACTGATTGTGTGGACTTTACTGCTCT 155
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThr 60
DB 156 TGGTGTGGACATGCCGTTTCATTCACCAATTCCTGGTGAATGGCTAAGAGTTTACA 215
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTspAla 80
DB 216 AATGTCATTCCTTAAGTGGACGTGGAGCACTAAGTCTGTTGCTCAGATGGGCT 275
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
:::|||||

Db 276 GTTAGGCTATGCCAACATTGGTTGTGTAAGAGAGAACCAATTGGGCAAGTGTT 335
QY 101 GYAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAla 120
|||||:::|||||
Db 336 GGAGCAAGAAAGAGACTGCACAGACAATTGAGAGACATGTGGCTTCACTATGCT 395

RESULT 2

RCTHIORXN
LOCUS R.communis mRNA for thioredoxin. 603 bp mRNA linear PLN 04-APR-1996
DEFINITION R.communis mRNA for thioredoxin.
ACCESSION Z70677
VERSION Z70677.1 GI:125953
KEYWORDS thioredoxin.
SOURCE Ricinus communis (castor bean)
ORGANISM Ricinus communis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.

REFERENCE 1 (bases 1 to 603)
AUTHORS Szederkenyi,J., Dolgener,B. and Schobert,C.
TITLE cDNA expressed in Ricinus cotyledons
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 603)
AUTHORS Dolgener,B.
JOURNAL Direct Submission
TITLE Submitted (04-APR-1996) Dolgener E., University of Bayreuth,
Department of Plant Physiology, Universitaetsstr.30, Bayreuth,
Germany, 95440

FEATURES

Location/Qualifiers
source
1..603
/organism="Ricinus communis"
/mol_type="mRNA"
/cultivar="Sanguineus"
/db_xref="taxon:3988"
/clone="pEDRH01.8"
/tissue_type="cotyledon"
/clone_lib="Lambda Excll"
/dev_stage="Seedling"
23..379
/codon_start=1
/product="thioredoxin"
/protein_id="CAA94534.1"
/db_xref="GI:125953"
/translation="MAAEQGVIGCHTVEANNEQLQKNDTKGLVYVDFTSWGPGR
FIAPFLAELAKKLPNVTFLKVDVDELKTVAEWAVESMPTFELKEGKIMDKVVGAK
DELOQTIAKEMATAST"

BASE COUNT

179 a 101 c 157 g 166 t
ORIGIN

Alignment Scores:

Pred. No.: 1.01e-48 Length: 603
Score: 490.00 Matches: 89
Percent Similarity: 91.38% Conservative: 17
Best Local Similarity: 76.72% Mismatches: 10
Query Match: 79.67% Indels: 0
DB: 8 Gaps: 0

US-09-786-715-6 (1-120) x RCTHIOXN (1-603)

QY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
 Db 26 GCACGAGAAAGAGGCAAGTGTATGGGTGCCACACTGTGAGGCAATGGAAATGACAAATTG 85
 QY 24 GlnLysGlyAenGluSerLysLeuIleValValAspPheThrAlaSerTrpCysGly 43
 Db 86 CAGAGGGGAATGATACCAAGGGACTGATCGTGTGTGATTACTGCTTCAATGGTGGGA 145
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThrSerValIle 63
 Db 146 CCATGCGGTTTCATTCCTCTCTTCTTGGCTGAGCTGGCCAAAGAACTGCCAAATGTATCC 205
 QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
 Db 206 TTCCTGAGGTGGATGGATGAATGAACTGAAGACTGTGCTCAGGAGTGGGCTGTGGAGTCA 265
 QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
 Db 266 ATGCCAACCTTATGTTCTGCTGAAGGAGGGGAGATATGGACAAGTGGTGGAGCAAG 325
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer 119
 Db 326 AAGAGCAACTGCAGCAAACTATAGCGAACAACACATGGCTACTGCTCC 373

RESULT 3

AX505468
 LOCUS AX505468 345 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 163 from Patent WO0216655.
 ACCESSION AX505468
 VERSION AX505468.1 GI:23386705
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1
 REFERENCE
 AUTHORS Harper, J.F., Krops, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 JOURNAL Patent: WO 0216655-A 163 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)
 FEATURES
 source Location/Qualifiers
 1..345
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 BASE COUNT 90 a 68 c 90 g 97 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.65e-48 Length: 345
 Score: 482.00 Matches: 88

Percent Similarity: 89.38% Conservative: 13
 Best Local Similarity: 77.88% Mismatches: 12
 Query Match: 78.37% Indels: 0
 DB: 6 Gaps: 0

US-09-786-715-6 (1-120) x AX505468 (1-345)

QY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
 Db 4 GCTTCGAGAGAGGACCAAGTGTATGGGTGCCACACTGTGAGGCAATGGAAATGACAAATTG 63
 QY 24 GlnLysGlyAenGluSerLysLeuIleValValAspPheThrAlaSerTrpCysGly 43
 Db 64 CAGAGGGCTAATGATCCAAACTCTGTGGTGTGATTTCACGSGCTCTCTGGTGGGA 123
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThrSerValIle 63
 Db 124 CCATGCGGTTTCATTCCTCTCTTCTTGGCTGAGCTGGCCAAAGAACTGCCAAATGTATCC 183
 QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
 Db 184 TTCCTCAAGGTTGATACATGATGAATTCGAAGTGGTGGCAAGTGTATGGGCGATACAGGCG 243
 QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
 Db 244 ATGCCAACCTTTCATGTTTTCGAGGAGGGAAGATTTTGACAAAGTGTGGAGCCAAAG 303
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116
 Db 304 AAGATGAGCTTCAGCTACCATTCGCAACACTTGGCT 342

RESULT 4

AX088687
 LOCUS AX088687 480 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 9189 mRNA, complete sequence.
 ACCESSION AX088687
 VERSION AX088687.1 GI:21407461
 KEYWORDS FLI_CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 480)
 REFERENCE
 AUTHORS Haas, B.J., Volfovsky, N., Town, C.D., Trukhan, M., Alexandrov, N.,
 Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
 TITLE Full-length messenger RNA sequences greatly improve genome
 annotation
 JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
 MEDLINE 22088475
 PUBMED 12093376
 REFERENCE 2 (bases 1 to 480)
 AUTHORS Brover, V., Trukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
 Feldmann, K.
 TITLE Full-length cDNA from Arabidopsis thaliana
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 480)

AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source Location/Qualifiers

1..480 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="9189"

CDS 13..357 /codon_start=1 /product="thioredoxin h" /protein_id="AA067038.1" /db_xref="GI:21617958" /translation="MASEEQVIACHTETVWNEQLQKANEKTLIVVDFTSWGQPCR FTAFPEADLAKLPVLFVLDVDELKSVASDWAIOAMPTFMELKEGKILDKVVGAKK DELASTIAKHLA"

BASE COUNT 124 a 93 c 111 g 152 t

ORIGIN

Alignment Scores:

Pred. No.: 6.83e-48 Length: 480

Score: 482.00 Matches: 88

Percent Similarity: 89.38% Conservative: 13

Best Local Similarity: 77.88% Mismatches: 12

Query Match: 78.37% Indels: 0

DB: 8 Gaps: 0

US-09-786-715-6 (1-120) * AY088687 (1-480)

QY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
:::|||||

DB 16 GCTTCGGAAGAGGACAAAGTGATCCCTGCCACCCGTGACACATGGAGACGACGCTT 75
|||||

QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGly 43
|||||

DB 76 CAGAGCAAGTAATGATCCAAACTCTTGTGTGGTGTGATTCACGGCTTCTGGTGTGA 135
|||||

QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63
|||||

Db 136 CCAATGCTTCAATGCTCCATTCTTTGCTGATTGCTAAGAACTTCCTACGCTT 195
|||||

QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspIrpAlaIleGluAla 83
|||||

Db 196 TTCTCAAGGTGATCTAGTGAATGAAGTCGGTGGCAAGTGATTTGGCGATACGCGG 255
|||||

QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
|||||

Db 256 ATGCCACCTTCATGTTTGAAGGAGGAGAGATTTTGACAAAGTTGTGTGAGCAAG 315
|||||

QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116
|||||

Db 316 AAAGATGAGCTTCAGTCTACCATTTGCCAAACACTTGCT 354
|||||

RESULT 5

ATHIOARA

LOCUS ATHIOARA 497 bp mRNA linear PIN 16-JUN-1994

DEFINITION A.thaliana mRNA for thioredoxin H.

ACCESSION Z14084

VERSION Z14084.1 GI:16551

KEYWORDS thioredoxin H.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 497)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

AUTHORS Rivera-Madrid,R., Marinho,P., Brugidou,C., Chartier,Y. and Meyer,Y.

TITLE Nucleotide sequence of a cDNA clone encoding an Arabidopsis thaliana thioredoxin h

JOURNAL Plant Physiol. 102 (1), 327-328 (1993)

MEDLINE 94151431

PUBMED 8108503

REFERENCE 2 (bases 1 to 497)
Meyer,Y.

AUTHORS Direct Submission

TITLE Submitted (16-JUL-1992) Yves Meyer PhD, Laboratoire de Physiologie vegetale, Universite de Perpignan, Av de Villeneuve 52, Perpignan, 66860, FRANCE

FEATURES

source Location/Qualifiers

1..497 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /codon_start=1 /product="Thioredoxin H" /protein_id="CAA78462.1" /db_xref="GI:16552" /translation="MASEEQVIACHTETVWNEQLQKANEKTLIVVDFTSWGQPCR FTAFPEADLAKLPVLFVLDVDELKSVASDWAIOAMPTFMELKEGKILDKVVGAKK DELASTIAKHLA"

CDS 135 a 94 c 116 g 152 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:	7.11e-48	Length:	497
Score:	492.00	Matches:	88
Percent Similarity:	89.48%	Conservative:	13
Best Local Similarity:	77.68%	Mismatches:	12
Query Match:	76.37%	Indels:	0
PS:	8	Gaps:	0

COM protein - nucleic search, using frame plus p2n model

Run on: December 5, 2003, 10:27:40 ; Search time 251 Seconds
(without alignments)
1290.568 Million cell updates/sec

4	SerSerGluGluGlyGlnValIleSerCysHisrnrValGluGlnrPASPAspPheLysr	23
	::: ::: ::: ::: ::: ::: ::: :::	
D6	33 GCCTTCGGAAGAAGACACAGCTGATCGCTCCACACACGCTTGAGACATGGACAGCAGCCTT	92
	::: ::: ::: ::: ::: ::: ::: :::	
QY	24 GlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGly	43
	::: ::: ::: ::: ::: ::: ::: :::	
D6	93 CAGAGGGCTAATGAATCCAAACTCTTGTGTGGTGTGATTCTACGGCTTCCTGGTGTGGA	152
	::: ::: ::: ::: ::: ::: ::: :::	
QY	44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle	63
	::: ::: ::: ::: ::: ::: ::: :::	
D6	153 CCATGTCTTTTCATCGCTCCATCTCTTGTGATTGGCTAAGAAACTTTCCTAAGCTGCTT	212
	::: ::: ::: ::: ::: ::: ::: :::	
QY	64 PheIleLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla	83
	::: ::: ::: ::: ::: ::: ::: :::	
D6	213 TTCTCAAGGTGATGATCTGATGAATTGAATCGGTGGCAAGTGATTGGCGGATACAGGGC	272
	::: ::: ::: ::: ::: ::: ::: :::	
QY	84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlnValAlaLys	103
	::: ::: ::: ::: ::: ::: ::: :::	
D6	273 ATGCCACCTTCATGTTTGNAGGAGAGGAGATTTGGACAAAGTTGTTGGAGCCACAG	332
	::: ::: ::: ::: ::: ::: ::: :::	
QY	104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla	116
	::: ::: ::: ::: ::: ::: ::: :::	
D6	333 AAAGATGAGCTTCAGTCTACCAATGCCAAACACTTGCT	371
	::: ::: ::: ::: ::: ::: ::: :::	

Title: US-09-786-715-6
Perfect score: 615
Sequence: 1 MAGSSEEGOVISCHTEVNN..... GAKKDELOOKI O K H V A S A 120

Scoring table:	
BL0SUM62	
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Deop 6.0	Delext 7.0

Searched: 2552756 seqs. 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
```

```

-O -c/cyt2.1/USP70_spool_p/US9786715/runat_04122003.092646.3929/app_query.fasta_1.263
-DBN Geneseq -UNITS=bits -QENI=fastp -SUFFIX=pmr.ring -MINMATCH=C0.1 -LOOFCU=0
-LOOFEQ=1 -LOOFEQ2=1 -START1=- -END1=- -MAIRX=blsnum2 -TRANS=H0.4cdi
-LIST=45 -DOLALIG=200 -THR_SCORES=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMGT=20 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9786715.@CN1.1.396.@runat_04122003.092646.3929 -NCFU=6 -ICFU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=1.20 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=0.5 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=10 -YGAPEXT=7 -DELOP=6 -DELXT=7

```

Database : N Geneseq 19Jun03:*

1: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/cgdata/geneseq/geneseq-emb1/NA1992.DAT.*

Search completed: December 5, 2003, 12:26:09

Job time : 1946 secs

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	615	100.0	738	21	AAZ51739 Glycine max thior
2	485	78.9	3888	24	ABSS3097 DNA encoding thior
3	485	78.9	3888	24	ABN9581 Phaseolin promoter
4	482	78.4	345	24	ABZ12358 Arabidopsis thalia
5	482	78.4	345	24	ABN9587 Arabidopsis thalia
6	482	78.4	561	21	AAC31522 Arabidopsis thalia
7	482	78.4	563	21	AAC34121 Arabidopsis thalia
8	482	78.4	3129	24	ABSS3095 DNA encoding thior
9	482	78.4	3129	24	ABN9579 Phaseolin promoter
10	482	78.4	3888	24	ABSS3096 DNA encoding Oleos
11	482	78.4	3888	24	ABN9580 Phaseolin promoter
12	478	77.7	470	21	AAC37781 Arabidopsis thalia
13	477	77.6	4935	24	ABN9586 Promoter-Oleosin t
14	472	76.7	509	22	AAH87768 Peppermint plant o
15	445	72.4	653	16	AAQ99783 Plant SAR gene pl.
16	445	72.4	653	20	AAV62799 Tobacco SAR CHX in
17	445	72.4	653	20	AAV81683 Tobacco protein-sy
18	438	71.2	402	25	ABX21664 Human GDP-mannose
19	430	69.9	574	21	AAC31738 Catalpa speciosa t
20	429	69.8	560	21	AAC41961 Arabidopsis thalia
21	424	68.9	392	25	ABX18045 Human GDP-mannose
22	410	66.7	601	21	AAZ51740 Glycine max thior
23	408	66.3	326	21	AAZ31097 Plant microsateili
24	403	65.5	346	21	AAZ31096 Plant microsateili
25	398	64.7	320	21	AAZ31785 Plant microsateili
26	392	63.7	614	21	AAZ51741 Vernonia mespilifo
27	391	63.6	390	25	ABX19403 Human GDP-mannose
28	388	63.1	686	15	AAQ78205 Gene coding for pr
29	388	63.1	687	21	AAC66375 Rice thiorredoxin h
30	385.5	62.7	590	21	AAC38792 Arabidopsis thalia
31	381.5	62.0	360	24	ABZ12359 Arabidopsis thalia
32	376.5	61.2	357	24	ABZ13931 Arabidopsis thalia
33	376.5	61.2	480	21	AAC36542 Arabidopsis thalia
34	376.5	61.2	524	21	AAC33829 Arabidopsis thalia

35 376.5 61.2 563 21 AAC34211 Arabidopsis thalia
 36 376.5 61.2 452 21 AAC48656 Arabidopsis thalia
 37 376 61.1 292 21 AAA31978 Plant microsateili
 38 368 59.8 393 17 AAT10451 Hard wheat thior
 39 368 59.8 393 21 AAC62457 Wheat thiorredoxin
 40 367 59.7 382 21 AAC62456 Wheat thiorredoxin
 41 367 59.7 384 17 AAT10450 Soft wheat thior
 42 367 59.7 419 21 AAA31942 Plant microsateili
 43 367 59.7 572 21 AAC52069 Arabidopsis thalia
 44 366 59.5 870 21 AAZ51737 Monardica charanti
 45 364.5 59.3 576 25 ABX56868 Arabidopsis thalia

ALIGNMENTS

RESULT 1
 AAZ51739

ID AAZ51739 standard; cDNA; 738 BP.

XX AC AAZ51739;

XX DT 04-JUL-2000 (first entry)

XX Glycine max thiorredoxin cDNA-1.

XX Glycine max thiorredoxin; clone sahlc.pk001.117; chimeric gene; soybean;
 transgenic plant; seed storage protein; allergenicity; ss.

XX Glycine max.

XX Key Location/Qualifiers
 CDS 68..430
 /*tag= a
 /product= "thiorredoxin"

XX WO20014239-A2.
 PD 16-MAR-2000.

XX 07-SEP-1999; 99WO-US20420.

XX 08-SEP-1998; 98US-0099501.

XX (DUPO) DU FONT DE NEMOURS & CO E I.
 PI Allen SM, Thorpe CJ, Lu AL;

XX WPI: 2000-256987/22.

XX P-PSDE; AAY70451.

XX New isolated polynucleotide encoding thiorredoxin polypeptide is useful
 for producing transgenic plants with an altered level of thiorredoxin -

XX Claim 3; Page 30; 33pp; English.

XX The present cDNA sequence encodes Glycine max thiorredoxin protein. The
 cDNA was derived from clone sahlc.pk001.117, which was isolated from a

CC cDNA library prepared from soybean tissue sprayed with authority
CC herbicide. Chimeric genes encoding all or a portion of the thiorodoxin
CC protein, in sense or antisense orientation are constructed, wherein
CC expression of the chimeric gene results in production of altered levels
CC of the thiorodoxin protein in a transformed host cell. Thiorodoxin is
CC involved in the disassembly of seed storage proteins during germination
CC by reducing S-S bonds and in the bread making process. Over expression of
CC thiorodoxin in cereals may reduce the allergenicity of any transgenic
CC protein engineered into cereal crops with high sulphydryl content.
XX
SQ Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;

Alignment Scores:
Pred. No.: 4.7e-73 Length: 738
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-786-715-6 (1-120) x AA251739 (1-738)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
DB 68 ATGGCTGGCTCATCGAAGAGGACAGATGATTCAGTCCACACCGTTGAAGATGGAC 127
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 128 GATCAACTCCAGAGGGACACGATCCAGAACTCATTTGTGGATTTTACTGCTCT 187
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
DB 188 TGGTGTGGACATGGCTTTCATTGCCACATTCCTGGCTGAGCTGGCTAAGAGATTCACA 247
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
DB 248 AGTGTGATATTCCTAAAGGTGGATGTGACGAATTAAGAGTGTGTTCTCAGATGGGCT 307
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 308 ATTGAGSCTATGCCACTTTTGTGTTTGTGAAGAGGGAACGCTTCGACAAAGTGTG 367
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 368 GGAGCAAGAGGATGAGCTGCAGCAGAAATACAGAAACATGTGGCTTCAGTACTGCT 427

Search completed: December 5, 2003, 11:53:37
Job time : 258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 12:26:17 ; Search time 259 Seconds
(without alignments)
1539.892 Million cell updates/sec

Title: US-09-786-715-6

Perfect score: 615

Sequence: 1 MAGSSEGGQVISCHTVEWN.....GAKKQELQKQIKQKVASASA 120

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp

-Q=/cgn2_1/USPTO.spool.p/US09786715/runat_04122003_092649_4034/app.query.fasta_1.263

-DB=Published Applications NA -QPM=fastap -SUFFIX=p2n.cn3p -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100

-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSZF=500 -MINLEN=0

-MAXLEN=200000000 -USER=US09786715 @CGN 1.1.85 @runat_04122003_092649_4034

-KCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:

2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:

6: /cgn2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:

Sequence 3641, Ap
Sequence 1059, Ap
Sequence 1145, Ap

43 222 36.1 418 14 US-10-040-862-3641
c 44 222 36.1 464 10 US-09-998-598-1059
45 222 36.1 479 11 US-09-535-459-1145

Search completed: December 5, 2003, 14:00:05
Job time : 266 secs

13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:**
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:**
15: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:**
16: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:**
17: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:**

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	485	78.9	3888	9	US-09-897-898-10		Sequence 10, Appl
2	485	78.9	3888	12	US-10-032-201B-19		Sequence 19, Appl
3	482	78.4	345	10	US-09-938-842A-163		Sequence 163, Appl
4	482	78.4	345	12	US-10-032-201B-38		Sequence 38, Appl
5	482	78.4	3129	9	US-09-897-898-5		Sequence 5, Appl
6	482	78.4	3129	12	US-10-032-201B-14		Sequence 14, Appl
7	482	78.4	3888	9	US-09-897-898-7		Sequence 7, Appl
8	482	78.4	3888	12	US-10-032-201B-16		Sequence 16, Appl
9	477	77.6	4935	12	US-10-032-201B-33		Sequence 33, Appl
10	438	71.2	402	10	US-09-878-574-3723		Sequence 3723, Ap
11	424	68.9	392	10	US-09-878-574-104		Sequence 104, App
12	391	63.6	390	10	US-09-878-574-1462		Sequence 1462, Ap
13	387	62.9	540	12	US-10-349-782-12		Sequence 12, Appl
14	381.5	62.0	360	10	US-09-938-842A-164		Sequence 164, App
15	376.5	61.2	357	10	US-09-938-842A-1736		Sequence 1736, Ap
16	376.5	61.2	357	14	US-10-323-362-1		Sequence 1, Appl
17	368	59.8	393	12	US-10-091-841-5		Sequence 5, Appl
18	367	59.7	382	12	US-10-091-841-3		Sequence 3, Appl
19	361	58.7	369	12	US-10-194-865-8		Sequence 8, Appl
20	361	58.7	369	12	US-10-091-841-1		Sequence 1, Appl
21	358	58.2	659	12	US-10-349-782-9		Sequence 9, Appl
22	316	51.4	277	10	US-09-878-574-8135		Sequence 8135, Ap
23	300	48.8	260	10	US-09-878-574-12941		Sequence 12941, A
24	293	47.6	370	10	US-09-878-574-3169		Sequence 3169, Ap
25	289	47.0	268	10	US-09-878-574-5995		Sequence 5995, Ap
26	285.5	46.4	402	10	US-09-938-842A-225		Sequence 225, App
27	285.5	46.4	613	9	US-09-770-149-855		Sequence 855, App
28	283	46.0	270	10	US-09-878-574-13445		Sequence 13445, A
29	279	45.4	272	10	US-09-878-574-12181		Sequence 12181, A
30	274	44.6	297	12	US-10-349-782-11		Sequence 11, Appl
31	273	44.4	2453	12	US-10-431-304-1		Sequence 1, Appl
32	272	44.2	456	10	US-09-924-035A-677		Sequence 677, Appl
33	256	41.6	250	9	US-09-923-876-1288		Sequence 1288, Ap
34	255	41.5	390	10	US-09-938-842A-2688		Sequence 2688, Ap
35	252	41.0	267	10	US-09-878-574-13137		Sequence 13137, A
36	237	38.5	261	10	US-09-878-574-7974		Sequence 7974, Ap
37	229	37.2	440	9	US-09-770-444-914		Sequence 914, App
38	226	36.7	357	10	US-09-938-842A-147		Sequence 147, App
39	222	36.1	301	9	US-09-615-349-76		Sequence 76, Appl
40	222	36.1	318	12	US-10-032-201B-46		Sequence 46, Appl
41	222	36.1	413	14	US-10-102-524-1396		Sequence 1396, Ap
42	222	36.1	418	10	US-09-796-692-3641		Sequence 3641, Ap

c

c

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 11:18:30 ; Search time 2055 Seconds
(without alignments)
1419.239 Million cell updates/sec

Title: US-09-786-715-6
Perfect score: 615
Sequence: 1 MAGSSEGGVTSCHTVEEN.....GAKKDELOQKIQHVASA 120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=cg2_1/USPTO_spool_p/US09786715/runat_04122003_092647_3948/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09786715 &CGN 1 1 2874 @runat_04122003_092647_3948 -NCFU=6 -ICFU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=1.0 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*

- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fur:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	615	100.0	459	12	BI699372 sac36c07.
2	615	100.0	495	13	BQ627882 sac65b02.
3	615	100.0	504	10	BG508580 sac74r08.
4	615	100.0	572	13	BQ630469 sac07f11.
5	615	100.0	562	13	BQ630383 sac06f10.
6	611	99.3	445	13	BQ578741 sac60h12.
c	7	600	97.6	14	CA937588 sac21g03.
	8	583	94.8	12	BE803058 sr31c05.y
9	582	94.6	522	13	BQ786610 sac70f04.
	10	581	94.5	452	9 AY988470 ad02f07.y
11	581	94.5	466	9	AW569018 s174e02.y
	12	577	93.8	613	12 BG838879 Gc01_06f0
c	13	566	92.0	627	14 CD401101 Gm_c42312
	14	556	90.4	687	14 CD038084 UTFF1002
15	533	86.7	570	14	CA901886 PCSA14170
	16	528	85.9	572	14 CA411141 E434_F -P
17	515	83.7	632	14	CA896875 PCFP05423
	18	512	83.3	445	10 BE61C138 sg62b12.y
19	512	83.3	512	14	CA794815 Cae_BL_18
	20	502	81.6	498	14 CA901885 PCSG11713
21	501	81.5	311	12	BI094898 sac61e08.
	22	500	81.3	537	13 BQ625213 USDA-PP_0
23	500	81.3	655	14	CB291687 UCRCS01_0
	24	497	80.8	653	14 CB293759 UCRCS01_0
25	495	80.5	517	14	CA795737 Cae_BL_27
	26	494	80.3	654	14 CB291686 UCRCS01_0
c	27	490	79.7	394	14 CA901887 PCSG15426
	28	490	79.7	493	14 CB303530 wsa0001_1
29	490	79.7	546	10	BE053835 GA_Ea000
	30	485	78.9	480	13 BUB22062 UB32BFB12

31 485 78.9 555 13 BU875659 BU875659 V009301 P
32 485 78.9 564 14 CA825633 CA825633 R63G02 tw
33 485 78.9 587 13 BU863412 BU863412 S027G05 P
34 485 78.9 622 9 AW255457 AW255457 M1480 pep
35 483 78.5 579 14 CA825259 CA825259 N58U03 tw
36 481 78.2 446 14 CA824794 CA824794 R48E09 tw
37 480 78.0 581 14 CA824034 CA824034 R35D03 tw
38 475.5 77.3 543 14 CB077020 CB077020 hj47G06.g
39 470 76.4 634 9 AU300651 AU300651 AU300651
40 469 76.3 586 9 AW255195 AW255195 M1185 pep
41 468 76.1 316 12 BI892660 BI892660 saf36C06.
42 468 76.1 413 14 T24347 T24347 crs1466 lam
43 467 75.9 540 14 CB377001 CB377001 HB02G03 L
44 467 75.9 602 10 BE054543 BE054543 GA_Ea002
45 466 75.8 523 12 BM111010 BM111010 EST558546

ALIGNMENTS

RESULT 1
BI699372 459 bp mRNA linear EST 29-NOV-2001
LOCUS seg36c07.yl Gm-cl081 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION ID: Gm-cl081-878 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; ,
mRNA sequence.

ACCESSION BI699372
VERSION BI699372.1 GI:15662001
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 459)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Streptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

TITLE Unpublished
JOURNAL Contact: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Putative full length read
vector to vector length is 527 This clone is available through:
ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
35801 For further information call: (800)-533-4363 or contact via
email: cdu@resgen.com
Seq primer: -403P from Gibco
High quality sequence stop: 424.

FEATURES
source

1..459
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl081-878"
/tissue_type="Roots of 7 day old 'Bragg' seedlings"
/dev_stages="7 days old"
/lab_host="DH1CB"
/clone_lib="Gm-cl081"
/note="Vector: pBluescript II SK+; Site1: EcoRI; Site2:
XhoI; The mRNA was isolated from roots of 7 day old
'Bragg' seedlings that were mock-infected 48 hours prior
to harvest. Dr. Gary Stacey generously donated the
tissue. The roots were flash-frozen in liquid nitrogen.
Stratagene's cDNA Synthesis Kit (catalog number 200401)
was used to synthesize the cDNA. First-strand synthesis
was performed with 5-methyl dCTP, hence the ligated cDNA
was hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor',
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAGAGAGAG(T)18V] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA, ligated to EcoRI adapters and subsequently
phosphorylated. The cDNA was then precipitated and
redissolved in sterile, RNase-, DNase-free water. The XhoI
site within the first- strand synthesis primer was then
restricted by digestion with XhoI from Promega (40U/ul);
all XhoI sites in the cDNA would be protected by their
hemimethylated status. The cDNA constructs were
size-fractionated with a 500bp cutoff, using Sephacryl
S-500 High Resolution (Pharmacia Biotech) in a 2-mm
diameter column and a bed volume of approximately 1ml. The
column eluent was precipitated, redissolved, and ligated
into Stratagene's pBluescript II XR Predigested vector
(pBluescript II SK(+)) vector that has been digested with
EcoRI and XhoI, and phosphorylated by Stratagene). This
library was constructed in the laboratory of Dr. Paul Keim
and Dr. Virginia H. Coryell at Northern Arizona
University."

BASE COUNT 132 a 81 c 114 g 132 t
ORIGIN

Alignment Scores:

Pred. No.: 6,61e-72 Length: 459
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DE: 12 Gaps: 0

US-09-786-715-6 (1-120) x BI699372 (1-459)

QY 1 MetaGlySerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsn 20
|||||
DB 33 ATGGCTGGCTCATCGAGAGAGGACCAAGTCATTAGCTGCCACACCGTGAAGTGAAC 92

vector to vector length is 605
Seq primer: -40RP from Gibco
High quality sequence stop: 420.

FEATURES
source
1..495
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl073-3915"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/clone_lib="Gm-cl073"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of *Fusarium solani* f.
sp. *glycines* (Plant Cell Report 18:375-380). Cultivar
Williams 82 is susceptible to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E. coli ElectroMax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Keena Philip and Steve Clough
(Lila Vodkin lab, University of Illinois)."

BASE COUNT 138 a 89 c 120 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 7.31e-72 Length: 495
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ627882 (1-495)

Qy 1 MetAlaGlySerSerGluGlyGlnValIleSerCysHisThrValGluIleTrpAsn 20
|||||
Db 35 ATGGCTGGCTCAICGGAGAGGGACAGTCATTAGTCGCCACCGCTTGAGATGGAC 94
|||||
Qy 21 AspGlnLeuGlnLysGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSer 40
|||||

21 AspGlnLeuGlnLysGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSer 40
|||||
Db 93 GATCACTCCAGAGGGCAACGATCCAGAACTCACTGTTGTGATTTTACTGCTCT 152
|||||
Qy 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
|||||
Db 153 TGGTGTGGACCATCGCTTCAATTCACCATCTCTTGGCTGAGCTGGCTAAGAGTCCAC 212
|||||
Qy 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
|||||
Db 213 AGTGCATATTCCTAAAGGTGGAGTGGAGCAATTAAGAGGTGTTCTCAAGATGGCT 272
|||||
Qy 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
|||||
Db 273 ATTGAGGCTATGCCACTTTTGTGTTGTGAAGAGGGAGCGCTCTCGACAAAGTGTG 332
|||||
Qy 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
Db 333 GGAGCAAGAAGATGAGTGGAGCGAGCAAAATACAGAAATGCTGCTCAGCTAGTGTCT 392
|||||

RESULT 2
BQ627882
LOCUS
DEFINITION
Gm-cl073-3915 5' similar to FR:Q43636 Q43636 THIOREDOXIN. ;, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 495)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cou@resgen.com web site:
www.resgen.com
Putative full length read


```

/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl072-1045"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/clone_lib="Gm-cl072"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar PI
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E. coli ElectroMax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Steve Clough (Lila Vodkin lab,
University of Illinois)."
BASE COUNT 143 a 123 g 150 t
ORIGIN
Alignment Scores:
Pred. No.: 7,49e-72 Length: 504
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-766-715-6 (1-120) x B3508580 (1-504)
QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
Db 37 ATGGCTGGCTCATCGGAGAGGAGCAAGTCATTAGCTGCCACACCGTTGAAGAATGGAAC 96
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
Db 97 GATCAACTCCAGAGGAGGCAACGAATCCAGAACTCATTTGTTGGATTTTACTGCTTCT 156
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
Db 157 TGGTGTGACCAATGCGGTTTCATTGCACCATTCITGCTGAGCTGGCTAGAGGTTCACA 216
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80

```

95 GATCAACTCCAGAGGAGGCAACGAATCCAGAACTCATTTGTTGGATTTTACTGCTTCT 154

41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60

155 TGGTGTGACCAATGCGGTTTCATTGCACCATTCITGCTGAGCTGGCTAGAGGTTCACA 214

61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80

215 ASGTGTCATATTCCTTAAGGTGGATGTGGAGAAATTAAGAGTGTTCCTCAAGATTGGCT 274

81 ILGLUAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100

275 ATTGAGGCTATGCGCACTTTTGTGTTTGAAGAGGGAACGCTTCGTGCAAAAGTGTG 334

101 GLYAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120

335 GGAGCGAAGNAGGATGAGCTGCACGAAATACAGAAACATGTGGCTTCAGCTAGTCT 394

RESULT 3

BG508580

LOCUS

DEFINITION

sac74f08.v1 Gm-cl072 Glycine max cDNA clone GENOME SYSTEMS CLONE

ID: Gm-cl072-1048 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE

1 (bases 1 to 504)

Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyle, J., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Korn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 256 1800

Fax: 314 256 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 439.

FEATURES

source

1..504

/organism="Glycine max"

/mol_type="mRNA"

/note="Vector: pBluescriptII SK+; Site 1: EcoRI; site 2: XhoI; This cDNA library was constructed from mRNA isolated from etiolated hypocotyl tissue of 9-10 day old seedlings of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dI) primer with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and XhoI. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 160 a 105 c 134 g 173 t
ORIGIN

Alignment Scores:
Pred. No.: 8.87e-72 Length: 572
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DE: 13 Gaps: 0

US-C9-786-715-6 (-120) x BQ630469 (1-572)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
DB 52 ATGGCTGGCTCATCGAGAGGGAACAGTATTAGTGGCCACACCGTTGAAGATGGAAC 111
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 112 GATCAACTCCAGAGGGAACGAATCCAGAACTCATTGTTGTGGATTTCCTGCTCT 171
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
DB 172 TGGTGTGACCACTGCGGCTTCATTGACCACTTCITGGCTGAGCTGGCTAAGAGTTCACA 231
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
DB 232 AGTGTCATATTCCTAAAGTGGATGTGGACGAATTAAGAGTGTTCCTCAAGATTGGCT 291
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 292 ATTTAGGCTATGCCCACTTTTGTCTTTGTGAAGAGGGAACGCTTCTGGACAAAGTGTG 351
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 352 GGAGCAAGAGAGATGAGCTGCACGACGAGAAATACAGAAACATGTGGCTTCAGCTAGTCT 411

RESULT 5
BQ630383
LOCUS
DEFINITION
saq06f10.y1 Gm-c1045 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1045-4100 5' similar to TR:Q43636 THIOREDOXIN. ; mRNA
sequence.
BQ630383
ACCESSION

Db 217 AGTGTCAATTCCTAAAGTGGATGTGGACGAATTAAGAGTGTTCCTCAAGATTGGCT 276
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 277 ATTTAGGCTATGCCCACTTTTGTCTTTGTGAAGAGGGAACGCTTCTGGACAAAGTGTG 336
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 337 GGAGCAAGAGAGATGAGCTGCACGAGAAATACAGAAACATGTGGCTTCAGCTAGTCT 396

RESULT 4
BQ630469
LOCUS
DEFINITION
saq07f11.y1 Gm-c1045 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1045-4125 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.
BQ630469
BQ630469.1 GI:21678108
EST.
KEYWORDS
SOURCE
Glycine max (soybean)
Glycine max
Glycine max
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 572)
Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Corvelli,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

TITLE
JOURNAL
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 433.
Location/Qualifiers
1. .572
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1045-4125"
/tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
/lab_host="DH10B"
/clone_lib="Gm-c1045"

FEATURES
source

VERSION BQ630383.1 GI:21678032
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 582)
 AUTHORS Shoemaker R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 296 1800
 Fax: 314 296 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40R2 from Gibco
 High quality sequence stop: 430.

FEATURES
 source
 1..582
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl045-4100"
 /tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
 /lab_host="DH10B"
 /clone_lib="Gm-cl045"
 /note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from etiolated hypocotyl tissue of 9-10 day old seedlings
 of the cultivar Williams 82. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) primer with a XhoI restriction site. EcoRI
 adapters were ligated to the blunt-ended cDNA fragments
 followed by digestion with EcoRI and XhoI. The cDNA
 fragments were directionally cloned into the EcoRI-XhoI
 restriction site of the pBluescript vector. The ligated
 cDNA fragments were transformed into DH10B host cells
 (Gibco BRL). This library was constructed by Dr. Randy
 Shoemaker."

BASE COUNT 166 a 105 c 135 g 176 t
 ORIGIN
 Alignment Scores: 9.08e-72 Length: 582
 Pred. No.:

Score: 615.00 Matches: 120
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-786-715-6 (1-120) x BQ630383 (1-582)

Qy 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
 Ds 52 ATGGCTGGCTCATCGGAGAGGGACCAAGTCATTAGCTGCCACCGCTGAGSAAATGGAAAC 111
 Qy 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeulleValValAspPheThrAlaSer 40
 Ds 112 GATCAACTCCAGAAAGGGCAAGCAATCCAAAGAACTCATTTGTTGTTGATTTTACTGCTTCT 171
 Qy 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
 Ds 172 TGGTGGGACCATGGCGTTTCATTGCACCATCTTGGCTGAGCTGGCTAAGAAAGTTCACA 231
 Qy 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
 Ds 232 AGTGTCATATTCCTAAAGGTGGATGTGGACGAATTAAGAGTGTTTCTCAAGATTGGGCT 291
 Qy 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
 Ds 292 ATTGAGGCTATGGCCCACTTTTGTGTTGTGAAGAGGGAACCGCTTCGGCAAAAGTGTG 351
 Qy 101 GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
 Ds 352 GGAGCAAAAGAGGATGAGCTGCACGAGAAATACAGAAACATGTGGCTTCAGCTAGTCT 411

RESULT 6
 BQ578741
 LOCUS BQ578741 445 bp mRNA linear EST 17-SEP-2002
 DEFINITION sar60h12.Y1 Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl074-6911 5' similar to IR:Q43636 Q43636 THIOREDOXIN. ; mRNA
 sequence.
 ACCESSION BQ578741
 VERSION BQ578741.1 GI:23063968
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 445)
 AUTHORS Shoemaker R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished
 COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: cu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 409.

FEATURES

source
 1. .445
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl074-6911"
 /tissue_type="seedlings induced for HR (hypersensitive response)"
 /dev_stage="9-11 day old"
 /lab_host="DH103"
 /clone_lib="Gm-cl074"
 /note="Vector: pBluescript II SK+; Site1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 130 a 81 c 118 g 116 t
 ORIGIN

Alignment Scores:

Pred. Nc.: 2.18e-71 Length: 445
 Score: 611.00 Matches: 119
 Percent Similarity: 99.17% Conservative: 0
 Best Local Similarity: 99.17% Mismatches: 1
 Query Match: 99.35% Indels: 0
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x EU578741 (1-445)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 79 ATGGCTGGATCATCGGAGAGGGGCAAGTCATTAGCTGCACACCGTTGAGAAATGGAAC 138
 QY 21 AspGlnLeuGlnLysGlyAsrGluSerLysLysLeuIleValValAspPheThrAlaSer 40
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 139 GATCAACTCCAGAGGCAAGCAAGTCCAGAACTCATTTGTGTGGATTTCAGCTCTCT 198
 QY 41 TrpCysGlyProCysArgPheLeuAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 199 TGGTGGGACCATGGCGTTCTTTGCACCATCTTGGCTGAGCTGGCTAAGAACTTACA 258
 QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 259 AATGTCTATTCTCTAAAGGTGGATGTGGACGAATTAAGAGTGTCTCTCAGATTGGCT 318
 QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 319 ATGAGGCTATGGCCACTTTTGTGTTGAAAGAGGGAACGCTTCGACAAAGTGTGTG 378
 QY 101 GValaLysLysAspGluLeuGlnLysLysLysLysHisValAlaSerAla 120
 D6 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 379 GGAGCAAGAAAGGATGAGCTCCAGCAAAATACAGAAACATGTGGCTTCAGCTATGCT 438

RESULT 7

CA937588/c
 LOCUS CA937588
 DEFINITION sav21g03.yl Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl048-6078 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
 sequence.
 CA937588 572 bp mRNA linear EST 30-DEC-2002
 VERSION CA937588.1 GI:27426068
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE

AUTHORS
 1 (bases 1 to 572)
 Shoenaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
 ,A., Boll,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project

TITLE

JOURNAL
 COMMENT
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

FEATURES

source

1..572
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl048-6078"
/tissue_type="whole seedlings of greenhouse grown plants"
/dev_stage="1 week old"
/lab_host="DH10B"
/clone_lib="Gm-cl048"
/note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 1 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Reim's laboratory at Northern
Arizona University."
BASE COUNT 174 a 125 c 102 g 171 t
ORIGIN
Alignment Scores:
Pred. No.: 9.03e-70 Length: 572
Score: 600.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.56% Indels: 0
DB: 14 Gaps: 0
US-09-786-715-6 (1-20) x CA937588 (1-572)
QY 4 Sex:Ser:Glu:Gly:Gln:Val:Ile:Ser:Cys:His:Thr:Val:Glu:Glu:Trp:Asn:Asp:Gln:Leu 23
|||||
DB 572 TCATCGNAGAGGACACAGTCATTAGCTGCCACACCGTTGAGATGGAGATCACTC 513
QY 24 Gln:Lys:Gly:Asn:Glu:Ser:Lys:Lys:Leu:Leu:Val:Val:Asp:Phe:Thr:Ala:Ser:Trp:Cys:Gly 43
|||||
DB 512 CAGAAAGGCGACGAAATCCAGAAACTCATTGTTGGATTCTTCTCTTGGTGGG 453
QY 44 Pro:Cys:Arg:Phe:Ile:Ala:Pro:Phe:Leu:Ala:Glu:Leu:Ala:Lys:Lys:Phe:Thr:Ser:Val:Ile 63
|||||
DB 432 CCATGGCGTTTCATGCACCATCTCTTGGCTGAGCTGGGTGAGAGTCCACAGTGTCATA 393
QY 64 Phe:Leu:Lys:Val:Asp:Val:Asp:Glu:Leu:Lys:Ser:Val:Ser:Gln:Asp:Trp:Ala:Ile:Glu:Ala 83

|||||
DB 392 TTCCTAAAGGTGATGTGGACGAATTAAAGAGTGTCTCTCAAGATTGGGCTATTGAGCT 333
QY 84 Met:Pro:Thr:Phe:Val:Phe:Val:Lys:Glu:Gly:Thr:Leu:Leu:Asp:Lys:Val:Val:Gly:Ala:Lys 103
|||||
DB 352 ATGCCCACTTTTGTGTTTGTGAAGAGGAGGACGCTCTTGACAAAGTGTGTGGAGCAAG 273
QY 104 Lys:Asp:Glu:Leu:Gln:Gln:Lys:Ile:Gln:Lys:His:Val:Ala:Ser:Ala:Ser:Ala 120
|||||
DB 272 AAGCATGAGCTGCAGCAGAAATAACAGAAACATGTGGCTTCACGCTAGTGCT 222

RESULT 8

BE803058

LOCUS

DEFINITION

st:cl051.y1 Gm-cl051 Glycine max cDNA clone

Gm-cl051-1569 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA

sequence.

BE803058

VERSION

BE803058.1 GI:10234170

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 472)

Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna

A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

Y., Person,B., Swaller,T., Gibbons,M., Fape,D., Harvey,N., Schurk

R., Ritter,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

R., Waterston,R. and Wilsch,R.

Public Soybean EST Project

Unpublished

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63105, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Putative full length read

vector to vector length is 622 This clone is available through:

ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL

35801 For further information call: (800)-533-4363 or contact via

email: ccu@resgen.com

High quality sequence stop: 441.

location/Qualifiers

1..472

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-1569"

/tissue_type="floral meristematic mRNA"

/lab_host="DH10B"

/clone_lib="Gm-cl051"

/notes="Vector: pBluescript II SK+; Site₁: EcoRI; Site₂: XhoI; The cDNA library was constructed from floral meristematic mRNA provided by Dr. Halina Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Schoemaker."

132 a 98 c 116 g 134 t 2 others

BASE COUNT
ORIGIN

Alignment Scores:		
Pred. No.:	1.32e-67	472
Score:	583.00	118
Percent Similarity:	99.16%	Conservative: 0
Best Local Similarity:	99.16%	Mismatches: 1
Query Match:	94.80%	Indels: 1
DE:	10	Gaps: 0

US-09-786-715-6 (1-120) x BE803058 (1-472)

2 AlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTyrAsnAsp 21
15 GCTGGCTCATCGGAAGAGGGACAAGTCATTAGCTGCCACACCGTTGAAGNATGGAAAGAT 74

Qy 22 GlnLeuGlnLysGlyAsnGluSerLysLysLeuValValAspPheThrAlaSerTrp 41
|||
Db 75 CAACTCCAGAGGGCAAGAAATCCAGAAGACTCATGTGTGGATTTTACTGCTTCTGG 134

42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61
135 TGTGGACCATGGCGTTTCATTGGACCATCTTGGCTGAGCTGGCTAGAGTTTCACAGT 194

Qy 62 ValIlePheLeuIysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIle 81
 |||||
 Db 195 GTCAATATCCTAAAGGT-GATGTGGAGCAATTAAAGAGTGTTTTCTCAAGATTGGGCTATT 253

Qy	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

QY 102 AlaLysLysAspGluLeuGlnGlnLysLysLysHisValAlaSerAla 120
 |||||
 Db 314 GCAAAGAGGATGAGCTGCAGCAGAAAAATACAGAAACATCTGGCTTCAGCTATGCT 370

RESULT 9			
BQ786610			
LOCUS	BQ786610	522 bp	mRNA linear EST 26-JUL-2002

Accession: BQ786610
 Sequence: Gm-c1076-5240 5' similar to TR:Q43636 Q43636 THIOREDUXIN. ; mRNA

R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 413.

TITLE
JOURNAL
COMMENT

BASE COUNT 142 a 97 c 120 g 163 t
ORIGIN

Alignment Scores:
Pred. No.: 2.05e-67 Length: 522
Score: 582.00 Matches: 113
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.63% Indels: 0
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ786610 (1-522)

Qy 8 GylGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 27
|||||
Db 1 GGACAGCATTAGCTGCCACACCGTTGAAGAATGGAACGATCACTCAAGAGGCAAC 60
|||||

Qy 28 GluSerLysLysLeulleValValAspPheThrAlaSerTrpCysGlyProCysArgPhe 47
|||||
Db 61 GAATCCAGAAACTCATTGTTGTGATTTTACTGCTTCTTGGTGGACCATGCCGTTTC 120
|||||

Qy 48 IleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLysVal 67
|||||
Db 121 ATGCGACCATCTTGGCTGAGCTGGCTAAGAGTTCAGAGTTCAGATTCCTTAAGGTG 180
|||||

Qy 68 AspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThrPhe 87
|||||
Db 181 GATGTGGACGAATTAAAGAGTGTCTCAAGATTGGGCTATTGAGGCTATGCCACTTTT 240
|||||

Qy 88 ValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysLysAspGluLeu 107
|||||
Db 241 GTCTTTGTAAAGGGAACGCTTCTGGACAAATGCTGGGAGCAAGAAAGGATGAGCTG 300
|||||

Qy 108 GlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
Db 301 CACAGAAATACAGAACATGTGGCTTCAGCTAGTGGT 339
|||||

RESULT 10
AI988470

LOCUS
AI988470 452 bp mRNA linear EST 30-NOV-2001
DEFINITION
sd02f07.y1 Gm-cl020 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl020-542 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.

ACCESSION
AI988470

VERSION
AI988470.1 GI:5820264

KEYWORDS
EST.

SOURCE
Glycine max (soybean)

ORGANISM
Glycine max

REFERENCE
1 (bases 1 to 452)
Shoemaker,R., Keim,P., Vockin,L., Erpelting,J., Corvett,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Thelasing,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

FEATURES
source
1..452
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl020-542"
/tissue_type="root nodules of greenhouse grown plants"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl020"
/notes="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from nodules on the roots of 2.5 month-old Glycine max
'Williams' plants that were greenhouse grown. The cDNA
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. First strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor'
nucleotide (v=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGACTAGTCTCGAG(T)18V] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA polymerase and size- fractionated with a
400 bp cutoff, using a sizeSep 400 Spun column from
Pharmacia. The column eluent was ligated to EcoRI adaptors
and phosphorylated. The XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size- fractionated with a 500 bp cutoff,
using GibcoBRL life Technologies' cDNA Size Predigested
vector (pBluescript II SK+) that has been digested with
EcoRI and XhoI, and phosphorylated by Stratagene. Both the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=56) and
sequence (n=16). This library was constructed by Dr. Paul
Keim and Dr. Virginia Corvett."

BASE COUNT 136 a 80 c 113 g 123 t
ORIGIN

Alignment Scores:
Pred. No.: 2.3e-67 Length: 452
Score: 581.00 Matches: 112

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@eatson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35901 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 416.

FEATURES

1..466
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl031-435"
/tissue_type="Williams seedlings, minus the cotyledons"
/lab_host="DH10B"
/clone_lib="Gm-cl031"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from whole 'Williams' seedlings, minus the cotyledons, which were propagated on paper towels with distilled water for 5 days, incubated at 40 degrees C for 1 hour. The cotyledons were removed and the remaining tissue was flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA size fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=13 and 5, respectively). This library was constructed by Dr. Paul Kelm and Dr. Virginia Corvelli."

EASE COUNT 142 a 82 c 116 g 126 t
ORIGIN

Alignment Scores: 2.39e-67 Length: 466
Pred. No.:

Percent Similarity: 97.50% Conservative: 5
Best Local Similarity: 93.33% Mismatches: 3
Query Match: 94.47% Indels: 0
DB: 9 Gaps: 0

US-09-786-715-6 (1-120) * AT988470 (1-452)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
DB 55 ATGGCAGGCTCTCGGAGGAGGACAGATAGTGTGCACACCGTTGACGATGGAC 114
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeulleValValAspPheThrAlaSer 40
DB 115 GATCACTCCAGAGGCAACCAATCCAGAACTCAITTTGTGTGATTTTACTGCTCT 174
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
DB 175 TGGTGTGGACCATGCCGTTTCATTGCACCATTTCTGGCTGAGCTGGCTAGAAAGTTCACA 234
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
DB 235 AGTGCTGATTCCTAAAGTGGATGTGGAGAAATTAAGAGTGTTCCTCAAGACTGGCT 294
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 295 ATTGAGGCTATGCCCACTTTTGTGTTTGTGAAGAGGGAACACTTCTAAGCAAGTGGTG 354
QY 101 GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 335 GGAGCAAGAGAGATGAGCTGCAGCAGACATAGAGAAATATGTGGCTTACGTAGTCT 414

RESULT 11
AW569018 466 bp mRNA linear EST 03-DEC-2001
LOCUS sl74e02.y1 Gm-cl031 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl031-435 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.

ACCESSION AW569018
VERSION AW569018.1 GI:7233673
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 466)
AUTHORS Shoemaker,R., Keim,P., Vodka,L., Erpelting,J., Corvelli,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
X., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Willson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project

Score: 581.00 Matches: 112
Percent Similarity: 97.50% Conservatives: 5
Best Local Similarity: 93.33% Mismatches: 3
Query Match: 94.47% Indels: 0
DB: 9 Gaps: 0

US-09-786-715-6 (1-120) x AW569018 (1-466)

Qy 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsp 20
Dy 69 ATGGCAGGCTCTCGGAGGAGGACCAAGTATTAGCTGTACACCGTTGACGCAITGGAC 128
Qy 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeulleValValAspPheThrAlaSer 40
Dy 129 GATCAACTCCAGAGGAGCAACCAATCCAGAACTCATTTGTTGGATTTTACTGCTTCT 188
Qy 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
Dy 189 TGTGTGGACCATGCCGTTTCATTGACCATTTCTTGGCTGAGCTGGCTAAGCAAGTGCAC 248
Qy 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
Dy 249 AGTGTGATATCTTAAGCGTGAATGGAGCAATTAAGAGTGTCTTCTCAAGACTGGCT 308
Qy 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
Dy 309 ATTGAGGCTATGCCACTTTTGTGTTTGAAGAGGAGCAACTTCTAAGCAAGTGGTG 368
Qy 101 GlyAlaLysLysAspGluLeuGlnLysLysLeulleGlnLysHisValAlaSerAlaSerAla 120
Dy 369 GGAGCAAGAGGATGAGCTGCAGCAACAATAGAGAAATATGTGGCTTAGCTAGTGTCT 428

RESULT 12

BG838879/c 613 bp mRNA linear EST 25-MAY-2001
LOCUS Gc01_06f08_R Gc02_AFC_EOORC_cold_stressed_Glycine_clandestina SSH
DEFINITION Glycine clandestina cDNA clone Gc01_06f08, mRNA sequence.

ACCESSION BG838879
VERSION BG838879.1 GI:14205201
KEYWORDS EST.
SOURCE Glycine clandestina
ORGANISM Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 613)
AUTHORS Singh,J.A., Savoie,A.J., Chapados,J., Couroux,P., De Moors,A.,
Harris,L.J., Hattori,J., Ouellet,J., Robert,L.S., Sprott,D. and
Tinker,N.A.
TITLE Differentially Expressed Sequence Tags from Cold-Stressed Glycine
clandestina Seedlings

JOURNAL Unpublished
COMMENT Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA

OC6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@agr.ca.
Location/Qualifiers

FEATURES

source

1..613
/organism="Glycine clandestina"
/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_06f08"
/issue_type="Leaves, stem"
/clone_lib="Gc02_AFC_EOORC_cold_stressed_Glycine_clandestina SSH"
/notes="Vector: pGEM-T easy; Site 1: EcoRI; Plants incubated at 2 C under 12 hours light/day and harvested after 3 days of cold treatment. SSH (Subtraction Suppressive Hybridization) library made using Clontech PCR-Select cDNA subtraction kit. PCR amplicons were ligated into the pGEM-Teasy vector (Promega)."

BASE COUNT 191 a 145 c 108 g 164 t

ORIGIN

Alignment Scores:
Pred. No.: 1..18e-66 Length: 613
Score: 577.00 Matches: 113
Percent Similarity: 95.80% Conservatives: 1
Best Local Similarity: 94.96% Mismatches: 5
Query Match: 93.82% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-6 (1-120) x BG838879 (1-613)

Qy 2 AlaGlySerSerGluGlyGlnValIleSerCysHisThrValGluGluTrpAsp 21
Dy 553 GCTGGCTCATCGRAGAGGACAGTATTAGCTGCCACCGTTGAGCATGGAACGAT 494
Qy 22 GlnLeuGlnLysGlyAsnGluSerLysLysLeulleValValAspPheThrAlaSerTrp 41
Dy 493 CAACCTCCAAAGGGCAACGAATCCAGAACTCATTTGTTGGATTTTACTGCTTCTGG 434
Qy 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGlnLeuAlaLysLysPheThrSer 61
Dy 433 TGTGGACCATGCCGTTTTCATTGCAACCATTTCTTGGCTGAGCTGGCTAAGAGTTCTAGT 374
Qy 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIle 81
Dy 373 GTCAATATCTTAAAGGKGATGTGGACGAATTAAAGAGTGTTCCTCAAGATTGGGCTATT 314
Qy 82 GluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGly 101
Dy 313 GAGGCTATGCCACTTTTGTGTTTGTGAAGAGGAGCAACGCTCTCTGGCAAGTGTGGGA 254
Qy 102 AlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
Dy 253 GCAAGAGAGGATGAGCTGCGACGACACATAGAGAAACATGTGGCTTACGCTAGTGTCT 197

RESULT 13
 CD401101/c
 LOCUS 627 bp mRNA linear EST 07-JUN-2003
 DEFINITION Gm_ck2312 Soybean induced by Salicylic Acid Glycine max cDNA 3', mRNA sequence.
 ACCESSION CD401101
 VERSION CD401101.1 GI:31459073
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurossids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 627)
 AUTHORS Tian,A.-G., Wang,J., Cui,P., Han,X.-J., Xu,H., Cong,L.-J., Huang,X.-G., Wang,X.-L., Jiao,Y.-Z., Wang,B.-J., Wang,Y.-J., Zhang,J.-S., Chen,S.-Y. and Yu,J.
 TITLE Soybean Expressed Sequence Tags Sequencing
 JOURNAL Unpublished
 COMMENT Contact: Chen S-Y
 Plant Biotechnology Laboratory
 Institute of Genetics and Developmental Biology, CAS, China
 Datun road, Beijing 100101, China
 Tel: 86-10-64886859
 Fax: 86-10-64873428
 Email: sychen@genetics.ac.cn
 Email: sychen@genetics.ac.cn
 Seq primer: T7 primer.
 Location/Qualifiers
 1..627
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Kefeng 1"
 /db_xref="taxon:3847"
 /tissue_type="Seedlings"
 /dev_stage="two-week seedlings"
 /lab_host="XLI-Blue MRF" strain"
 /clone_lib="Soybean induced by Salicylic Acid"
 /note="Vector: pBluescript SK+; Site 1: EcoR I; Site 2: Xho I; The cDNA library was constructed by He, C-Y from mRNA isolated from two-week seedlings (cultivar Kefeng 1) treated by spraying 2.0mM salicylic acid for 24, 36, 48 and 72 h. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI-Blue MRF host cells (Stratagene)."
 BASE COUNT 178 a 135 c 111 g 203 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.62e-65 Length: 627
 Score: 566.00 Matches: 118

Percent Similarity: 80.27% Conservative: 0
 Best Local Similarity: 80.27% Mismatches: 0
 Query Match: 92.03% Indels: 29
 DB: 14 Gaps: 1
 US-09-786-715-6 (1-120) x CD401101 (1-627)
 Qy 3 GlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGln 22
 D6 621 GGTCTCATCGGAGAGGGGACAAAGTCATTAGCTGCCACACCGTTGAAGAATGGAAGCA 562
 Qy 23 LeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrpCys 42
 D6 561 CTCCAGAGGGCAACGAATCCAGAACTCATTTGTTGGATTTTACTGCTTCTGGTGT 502
 Qy 43 GlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerVal 62
 D6 501 GGACATGCCGTTTCAATTGCCACATCTCTGGCTGAGCTGGCTAAGAAAGTTCAGATGTC 442
 Qy 63 IlePheLeuLysValAspValAspGluLeuLys----- 73
 D6 441 ATATTCCTAAAGCTGGATGTGGACGAATTAAGAGTAAAAAACCCTACCAAAAAGCA 382
 Qy 74 -----Se 74
 D6 361 GAGAACTAAGAAAGAAATTCATTAGAAAGTAATCAAGATAATTTGTGTGATTCAGAG 322
 Qy 74 rValSerGlnAspTrpAlaIleGluAlaMetProThrPheValPheValLysGluGlyTh 94
 D6 321 TGTCTCTCAAGATTGGCTATTGGGCTATGCCACCTTTTGTGTTTGTGAAGAGGGAAC 262
 Qy 94 rLeuLeuAspLysValValGlyAlaLysLysAspGluLeuGlnLysLysLysLysHi 114
 D6 261 GCTTCTGGACAAAAGTGTGGGAGCAAGAGATGAGCTGCACAGAAATACAGAAACA 202
 Qy 114 sValAlaSerAlaSerAla 120
 D6 201 TGTGCTTCAGCTAGTGTCT 183
 RESULT 14
 CD038084
 LOCUS 687 bp mRNA linear EST 07-MAY-2003
 DEFINITION UPP1002_C04 USDA-Tifton Peanut Immature pod cDNA Library (UTP)
 ACCESSION CD038084
 VERSION CD038084.1 GI:30419922
 KEYWORDS EST.
 SOURCE Arachis hypogaea (peanut)
 ORGANISM Arachis hypogaea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurossids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
 REFERENCE 1 (bases 1 to 687)
 AUTHORS Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G. and Lynch,R.E.
 TITLE Generation and Analyses of ESTs for Arachis hypogaea

JOURNAL COMMENT

Unpublished
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 228-387-2334
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES

Location/Qualifiers
1..687
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="JTPP1002 C04"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="X11-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIZOL-Reagent ultrapur (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

BASE COUNT

189 a 107 c 176 g 215 t

ORIGIN

Alignment Scores:
Pred. No.: 8.92e-64 Length: 687
Score: 556.00 Matches: 103
Percent Similarity: 96.67% Conservative: 13
Best Local Similarity: 85.93% Mismatches: 4
Query Match: 90.41% Indels: 0
DB: 14 Gaps: 0

US-09-786-715-6 (1-120) x CD038084 (1-687)

QY 1 MetAlaGlySerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsn 20
|||||
DB 69 ATGGCTGCTTCATCAGAGAGGGACAGATTATCGGTTCCACTCGGTGATGATGGACC 128
|||||
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
|||||

Db 129 CAACAGCTCGAGAGGGAAATGAATCTAAGAAGCTGATTTGTTGATTTCACTGCTCA 188
|||||
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
|||||
Db 189 TGGTGGGACCATGCGGCTTCATTGCACCTTTTCTGGCTGAGTGGCTAAGAAGTTTACA 248
|||||
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
:::|||||
Db 249 AATGTCATATTTTGAAGTGGATGGATGAAGTCACTGCTGTTGCTCAAGACTGGGCT 308
:::|||||
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrIleLeuAspLysValVal 100
:::|||||
Db 309 GTGGAGCAATGCCACCTTTGTGTTTGTGAAGAGGAGCAACATTATTGCAAGTGTG 368
|||||
QY 101 GtValAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
Db 369 GGAGCAAGAGGAGGAGACTGCGAGCAAGAGCTTGAGAAACATGTGGCTACAGCCAGT 428
|||||

RESULT 15

CA901886 570 bp mRNA linear EST 27-DEC-2002
CA901886
LOCUS PCSC14170 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus
DEFINITION coccineus cDNA 5' similar to Thiofodoxin H-type, mRNA sequence.
ACCESSION CA901886
VERSION CA901886.1 GI:27388878
KEYWORDS EST.
SOURCE Phaseolus coccineus
ORGANISM Phaseolus coccineus

REFERENCE 1 (bases 1 to 570)
Bui,A.Q., Le,B.H., Weterings,K., Bi,Y.-P., Choi,J.-S., McElroy,K.E.,
AUTHORS / Choi,P.S., Harada,J.J., Fischer,R.L. and Goldberg,R.B.
TITLE Gene Activity in Different Regions of a Post-Fertilization Plant
EMBRYO by EST Analysis
JOURNAL Unpublished
COMMENT Contact: Goldberg, R.B.
Department of Molecular, Cell, & Developmental Biology
University of California, Los Angeles
621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA
Tel: 310 825 3270
Fax: 310 825 8201
Email: bobg@ucla.edu
Seq primer: 5' Triplex
POLYA-No.

FEATURES

Location/Qualifiers
1..570
/organism="Phaseolus coccineus"
/mol_type="mRNA"
/cultivar="Hammond's Dwarf Scarlet"
/db_xref="taxon:3886"
/dev_stage="6-days post-pollination"
/clone_lib="Scarlet Runner Bean Suspensor Region Triplex2"
/note="Organ: Suspensor Region of Globular-Stage Embryos;

Vector: Triplex2; Site 1: SfiI; Site 2: SfiI; Suspensor regions were micro-dissected from globular-stage embryos six days after pollination from greenhouse-grown plants [Weterings et al., Plant Cell 13, 2409-2425 (2001)]. Double-stranded cDNA was synthesized from suspensor mRNA using the SMART cDNA library construction kit according to the manufacturer (Clontech). The suspensor cDNA fragments were directionally ligated into the SfiI restriction site of the lambda Triplex2 vector (Clontech), and the recombinant cDNAs were transformed into E. coli XL1-Blue cells (Clontech). Suspensor cDNA plasmids used for directional sequencing were obtained by in vivo excision from the lambda Triplex2 recombinants in E. coli BM25.8 cells (Clontech)."

BASE COUNT 154 a 92 c 134 g 190 t
ORIGIN

Alignment Scores:

Pred. No.:	8.31e-61	Length:	570
Score:	533.00	Matches:	100
Percent Similarity:	95.00%	Conservative:	14
Best Local Similarity:	83.33%	Mismatches:	6
Query Match:	86.67%	Indels:	0
DB:	14	Gaps:	0

US-09-786-715-6 (1-120) x CA901896 (1-570)

QY	1	MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn	20
Db	82	ATGCGCGCGCATCAGAGAGGAGACAGTCATTCGCTGCCACACCGTTGAGGCATGGACA	141
QY	21	AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer	40
Db	142	GAACACATCGAAAGGCAATATTCACAGAACTAATGTGTGATTTACTGTCTCT	201
QY	41	TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThr	60
Db	202	TGGTGGGACCAATGCGGTTTCATTTCTCCATTTGGCTGAGCTGGCTAAGAGTTTACA	261
QY	61	SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla	80
Db	262	AATGCTATTTCGAAGTGGATGTAGACGATTAAAGAGTGTTCGAAGATTTTGTCT	321
QY	81	IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal	100
Db	322	GTGAGCAATCCCACTTTTGTGTGTAAGAGGGAATCTTCTGGCAAAAGTGTG	381
QY	101	GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla	120
Db	382	GGACAAAGAGGAGAGATTTGACGACACATAGAGAAACATGTGCTGCACCTAGTGTCT	441

Search completed: December 5, 2003, 13:00:41
Job time : 2064 secs